

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平8-297499

(43) 公開日 平成8年(1996)11月12日

(51) Int.Cl.<sup>6</sup>

G10L 3/02

識別記号

庁内整理番号

F I

G10L 3/02

技術表示箇所

A

審査請求 未請求 請求項の数 9 O L (全 9 頁)

(21) 出願番号 特願平7-102093

(22) 出願日 平成7年(1995)4月26日

(71) 出願人 000005821

松下電器産業株式会社

大阪府門真市大字門真1006番地

(72) 発明者 加藤 弓子

大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

(72) 発明者 吉住 嘉之

大阪府門真市大字門真1006番地 松下電器  
産業株式会社内

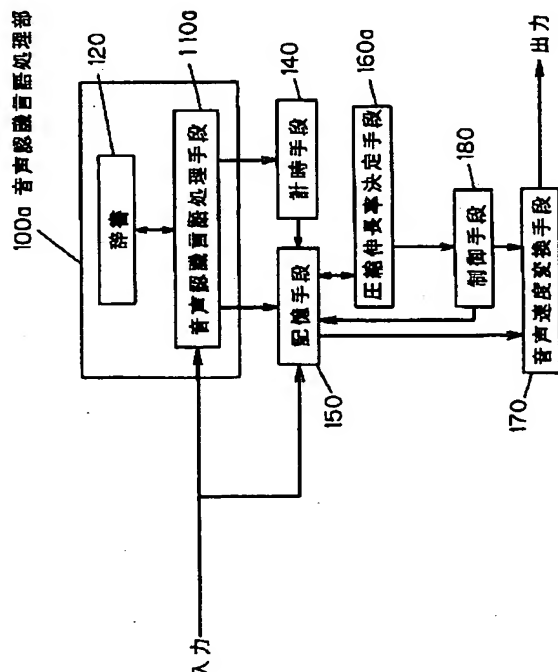
(74) 代理人 弁理士 滝本 智之 (外1名)

(54) 【発明の名称】 音声速度変換制御装置および音声速度変換制御方法

(57) 【要約】

【目的】 音声の出力時間の延長を抑制し、かつ話速の伸長の効果を拡大し、フレーズ後部の情報の欠落のない音声速度変換制御装置を提供することを目的とする。

【構成】 音声認識言語処理部100aと、音声認識言語処理部100aの出力にもとづき音声の時間長を処理単位ごとに計時する計時手段140と、時間長と前記音声認識言語処理部100aにより判断された重要度を処理単位ごとに記憶する記憶手段150と(数1)を用いて圧縮および伸長率を処理単位ごとに決定する圧縮伸長率決定手段160aと、前記記憶手段150から呼び出された入力信号を圧縮あるいは伸長する音声速度変換手段170と、前記圧縮伸長率決定手段160aの出力に基づき前記記憶手段150から音声を取り出し前記音声速度変換手段170を制御する制御手段180とを備える。



1

## 【特許請求の範囲】

【請求項 1】 入力された音声認識し言語的な意味処理を行い入力された音声を言語的情報量によって重要度を判断する音声認識言語処理部と、前記音声認識言語処理部の出力に基づき速度変換後の発話に変換前の発話と同じ時間長になるよう調整して各重要度に圧縮および伸長率を割り当てる時間長調整部と、前記時間長調整部の出力に基づき音声速度変換の圧縮および伸長率を制御する制御手段と、前記制御手段から出力される圧縮および伸長率で入力信号を圧縮あるいは伸長する音声速度変換手段とを備えた音声速度変換制御装置。

【請求項 2】 音声認識言語処理部は間投詞を冗長部分としてその冗長部分以外より低いあらかじめ定められた重要度を割り付け文の主部・述部を構成する品詞を情報量の多い部分としてその主部・述部を構成する品詞以外より高いあらかじめ定められた重要度を割り付ける請求項 1 記載の音声速度変換制御装置。

【請求項 3】 音声認識言語処理部は統計的に得られた発話データに基づいた音韻・単語・文法・構文・文脈・意味・対話の知識のうち少なくとも 1 つを制約として用いることによって構成された発話モデルに対する入力音声 \*

$i$ : 伸長優先順位 数字が小さいものほど優先順位が高い

$$1 \leq i \leq n$$

$$T_i = \Sigma (\text{伸長優先順位 } i \text{ の処理単位の時間長})$$

$$\Sigma T_i - T_1 \times x_1 = T_2 \times x_2 + T_3 \times x_3 + \dots + T_n \times x_n$$

$$\text{ただし } 0.3 \leq x_i \leq 3$$

$$x_1 > 1$$

$$x_1 > x_2 > x_3 > \dots > x_n$$

【請求項 8】 音声の処理単位ごとの圧縮あるいは伸長率を (数 2) を用いて決定する音声速度変換制御方法。 ※ 【数 2】

$i$ : 伸長優先順位 数字が小さいものほど優先順位が高い

$$1 \leq i \leq n$$

$$2 \leq k \leq 20 \quad k \text{ は定数}$$

$$T_i = \Sigma (\text{伸長優先順位 } i \text{ の処理単位の時間長})$$

$$\Sigma T_i - T_1 \times x_1 - \sum_{i=k}^n (T_i \times x_i) = \sum_{i=2}^{k-1} (T_i \times x_i)$$

$$\text{ただし } 0.3 \leq x_i \leq 3$$

$$i < k \text{ のとき } x_i > 1$$

$$i \geq k \text{ のとき } x_i \leq 1$$

$$x_1 > x_2 > x_3 > \dots > x_n$$

【請求項 9】 音声の処理単位ごとの圧縮あるいは伸長率を (数 3) を用いて決定する音声速度変換制御方法。 【数 3】

2

\* の適合度を言語的情報量の指標として重要度を判断する請求項 1 記載の音声速度変換制御装置。

【請求項 4】 音声認識言語処理部として音声認識言語処理手段と、辞書部とを備え、前記辞書部を参照して前記音声認識言語処理手段で入力音声の認識と意味処理を行う請求項 1 記載の音声速度変換制御装置。

【請求項 5】 音声認識言語処理部として音声認識言語処理手段と、複数の辞書部と、切り換え手段とを備え、前記切り換え手段によってあらかじめ接続を切り換えた前記辞書部を参照して前記音声認識言語処理手段で入力音声の認識と意味処理を行う請求項 1 記載の音声速度変換制御装置。

【請求項 6】 音声認識言語処理部として音声認識言語処理手段と、辞書読み取り手段とを備え、前記辞書読み取り手段によって記録媒体に格納された辞書を参照して前記音声認識言語処理手段で入力音声の認識と意味処理を行う請求項 1 記載の音声速度変換制御装置。

【請求項 7】 音声の処理単位ごとの圧縮あるいは伸長率を (数 1) を用いて決定する音声速度変換制御方法。

【数 1】

3

$i$ : 伸長優先順位 数字が小さいものほど優先順位が高い

$1 \leq i \leq L$   $L$ は定数

$T_i = \Sigma$  (伸長優先順位 $i$ の処理単位の時間長)

$\Sigma T_i = \Sigma(T_i \times x_i)$

$$x_i = r_i \times \frac{\Sigma T_i}{\Sigma(T_i \times r_i)}$$

ただし  $x_1 : x_2 : x_3 : \dots : x_L = r_1 : r_2 : r_3 : \dots : r_L$

$r_1, r_2, r_3, \dots, r_L$  は定数

$r_1 > r_2 > r_3 > \dots > r_n$

#### 【発明の詳細な説明】

##### 【0001】

【産業上の利用分野】本発明は、音声速度変換に関する信号処理分野における音声速度変換制御装置に関するものである。

##### 【0002】

【従来の技術】従来、音声の速度変換において入力音声を一定の倍率で伸長すると出力音声のデータ量は増えるので、全体の音声の出力時間は伸長倍率に応じて長くなる。また、画像や口唇の動きとのタイミングのずれが時間が経つにつれて次第に大きくなる。そこで上記のような画像や口唇の動きとのタイミングのずれが時間が経つにつれて増大するのを防ぐ方法として、音声速度変換の制御装置として次のようなものが報告されている(日本音響学会講演論文集I p.349 平成4年10月)。図7に従来の音声速度変換制御装置の構成ブロック図を示す。900は入力音声から無音区間を検出する無音区間検出手段、910は音声の有声音部分のピッチを抽出する有声部ピッチ抽出手段、920は無音区間終了直後から3つの有声区間のピッチを比較してPitch\_maxを決定し無音区間終了後から第 $n$ 有声区間のピッチとPitch\_maxを比較する判定部、180は速度変換の伸長、圧縮率を制御する制御手段、940は時間を計測する計時手段、170は音声の速度変換を行う音声速度変換手段である。

【0003】以上のように構成された従来の音声速度変換制御装置において無音区間検出手段900で無音区間が250msec以上の区間を検出し、図8の(1)で示したポーズとし、その直後に続く有声/無声の開始点をフレーズの開始点とする。有声部ピッチ抽出手段900で(2)で示したフレーズの開始点から3つの有声区間のピッチを抽出し、判定部920で第1から第3有声区間のうち最高ピッチ周波数をPitch\_maxとする。制御手段180で(3)で示した第1有声区間の開始点( $V_{st}$ )における話速の伸長倍率をと $r_s$ し、制御手段180に $r_s$ を伝達する。音声速度変換手段は制御手段180より伝達された伸長倍率で音声速度変換を行う。計時手段940は(4)で示した $V_{st}$ から、2000msecまでの時間を計測する。制御手段180により $V_{st}$ から、2000msec

4

までの範囲で話速の伸長倍率を $r_s$ から $r_e$ まで変化させ、制御手段180に伝達する。判定部920で(5)で示した第 $n$ 有声区間の( $n \geq k$ )における平均ピッチ周波数であるPitch( $n$ )とPitch\_maxを比較し、Pitch( $n$ )がPitch\_maxの70%より大きいときのみ制御手段180で第 $n$ 有声区間の伸長倍率を( $r_s - 0.1$ )とする。それ以外の時の伸長倍率は $r_e$ とする。制御手段180より伝達された伸長倍率により音声速度変換手段で(6)で示した $V2_{st}$ から、2000msecまでの範囲で、伸長倍率を( $r_s - 0.1$ )から $r_e$ まで変化させて音声速度変換を行う。次のポーズまで順次図8の(5)(6)に示した区間の動作を繰り返す。

##### 【0004】

【発明が解決しようとする課題】しかしながら上記のような構成では、発話内容の重要度に関わらず、直前のポーズからの時間的距離とピッチという物理特性によってのみ話速の伸長倍率を制御するため、フレーズの後部に重要な情報が含まれる場合には、話速の伸長の効果が得られないあるいは圧縮処理によって情報が欠落する可能性がある。

【0005】本発明は上記の従来の問題を解決するもので音声の認識と言語処理により発話内容の重要度を判断し発話中で重要度の高い部分の音声の伸長倍率を大きくし、重要度の低い部分の伸長倍率を小さくすることによって、音声の出力時間の延長を抑制し、かつ話速の伸長の効果を拡大し、フレーズ後部の情報の欠落のない音声速度変換制御装置を提供することを目的とする。

##### 【0006】

【課題を解決するための手段】上記の目的を達成するために本発明の音声速度変換制御装置における第1の発明においては、入力された音声認識し、意味的な言語処理により語の重要性を判断する音声認識言語処理部と、音声認識言語処理部の出力信号に基づき圧縮、伸長率の時間変化を決定する制御手段と、制御手段で圧縮、伸長率が決定される間の時間を補償する遅延手段と、前記制御手段から出力される圧縮、伸長率で入力信号を圧縮、伸長する音声速度変換手段とを備えた音声速度変換制御装置である。

【0007】また、本発明の音声速度変換制御装置にお

20

30

40

50

5

ける第2の発明においては、入力された音声認識し、意味的な言語処理により次に入力される語の重要性を予測する音声認識言語処理手段の出力信号に基づき圧縮、伸長率の時間変化を決定する制御手段と、前記制御手段から出力される圧縮、伸長率で入力信号を圧縮、伸長する音声速度変換手段とを備えた音声速度変換制御装置である。

#### 【0008】

【作用】上記した構成によれば、入力された音声認識し、発話中におけるそれぞれの語の重要性を判断し、重要性の高いものは伸長率を大きくし、重要性の低いものは伸長率を小さくするように音声速度変換手段を制御するため、話速の伸長の効果を拡大し、フレーズ後部の情報の欠落なく速度変換することができる。

#### 【0009】

##### 【実施例】

（実施例1）以下本発明の第1の実施例について、図面を参照しながら説明する。

【0010】図1は本発明の音声速度変換制御装置の第1の実施例を示す構成ブロック図である。図2にこの実施例におけるタイムチャートを示す。図1において図7と同一物または部分については同一符号を付しているので説明を省略し、異なった部分についてのみ説明する。図1において、100aは入力された音声認識し、言語的な意味処理を行う音声認識言語処理部であり、音声認識および意味処理による重要度の判定を行う音声認識言語処理手段110aと、音声認識と意味処理に用いる辞書部120からなる。180は音声認識言語処理部100aより伝達された音声の処理単位ごとの時間長を計時する計時手段であり、150は入力された音声と音声認識言語処理部100aより伝達された重要度と計時手段180より伝達された時間長を音声認識言語処理部100aより伝達された処理単位ごとに記憶する記憶手段であり、160aは記憶手段150に記憶された処理単位ごとの重要度と時間長より出力する音声の時間長が入力された音声の時間長に等しくなるように圧縮、伸長率を決定する圧縮伸長率決定手段である。170は記憶手段150から入力される信号の圧縮、伸長を行う音声速度変換手段であり、180は時間長調整部130aで決定された圧縮、伸長率に基づき、入力された音声処理単位ごとに記憶手段150より呼び出し、音声速度変換手段170で圧縮あるいは伸長するように記憶手段150と音声速度変換手段170を制御する制御手段である。

【0011】以上のように構成された本実施例の音声速度変換制御装置について、その動作を図1に従って説明する。入力音声は音声認識言語処理部100aと記憶手段150へ伝達される。音声認識言語処理部100aで音声認識言語処理手段110aは辞書部120を参照して音声認識し、認識した音声の意味的な言語処理を行

6

う。さらに音声認識言語処理手段110aは言語処理結果に基づき、処理単位が持つ情報量に応じて1からnまでの重要度を割り付ける。1フレーズ内でもっとも情報量が大きく重要な処理単位を重要度1とする。情報量の少ない冗長部として間投詞に他の品詞より大きな値の重要度を割り付け、情報量の多い部分として主部、述部を構成する品詞には重要度として小さな値を割り付ける。割り付けた重要度を、入力された音声の処理単位ごとに計時手段140および記憶手段150へ伝達する。計時手段140は音声認識言語処理手段110aより入力された処理単位ごとの時間長を測定し記憶手段150へ伝達する。記憶手段150は音声認識言語処理手段110aより伝達された処理単位ごとに入力音声、重要度および計時手段140より伝達された時間長を1フレーズ分記憶する。（数1）のiに音声認識言語処理手段110aで割り付けられた重要度をあてはめる。

【0012】図2の（a）に示すように、 $T_i$ を重要度iを割り付けられた処理単位の時間長の合計とする。圧縮伸長率決定手段160aは記憶手段150を参照し図2の（b）に示すように重要度1を割り付けられた処理単位を1.2倍に伸長する。1フレーズ全体の時間長すなわち1フレーズ内の全処理単位の時間長を合計した値を $\Sigma T_i$ とする。図2の（b）（c）に示すように、重要度iを割り付けられた処理単位に対する伸長あるいは圧縮率を $x_i$ とする。図2の（c）に示すように、重要度2からnを割り付けられた処理単位について $T_i$ に $x_i$ を乗じた値の合計が、 $\Sigma T_i$ から $T_1$ に1.2を乗じた値を減じたものと等しくなるよう各 $x_i$ を決定し、処理単位ごとの圧縮あるいは伸長率を制御手段180へ伝達する。ただしiが1以外の $x_i$ は0.8以上1.2未満であり、iがより大きい場合により小さい $x_i$ 、iがより小さい場合により大きい $x_i$ を与えるものとする。制御手段180は処理単位ごとの圧縮あるいは伸長率を音声速度変換手段170に伝達し、入力音声を記憶手段150より呼び出して音声速度変換手段170に伝達する。音声速度変換手段170は制御手段より伝達された圧縮あるいは伸長率に基づいて処理単位ごとに速度変換を行う。

【0013】（実施例2）次に本発明の第2の実施例について、図面を参照しながら説明する。

【0014】図3は本発明の第2の実施例における音声速度変換制御装置の部分を示す構成ブロック図である。図4にこの実施例のタイムチャートを示す。第2の実施例の構成において第1の実施例の構成と同一物または部分については説明を省略し、異なった部分についてのみ説明する。図1の音声認識言語処理部100aが、図3において音声認識言語処理部100bに置き変わり圧縮伸長率決定手段160aが160bに変わった以外は図1と同一な構成である。前記の音声認識言語処理部100bは音声認識言語処理手段110bと、あらかじめ設定された場面に応じた語彙を持つ辞書220aからd

7

と、音声認識言語処理手段 110b が参照する辞書 220a から d を場面に応じて切り替えるスイッチ 260 を有する。

【0015】 以上のように構成された本実施例の音声速度変換制御装置について、その動作を図 3 に従って説明する。まず第 1 の実施例と同様に入力音声は音声認識言語処理部 100b と記憶手段 150 へ伝達される。音声認識言語処理手段 110b はあらかじめ場面にに応じて切り替え入力により切り替えられたスイッチ 260 によって接続された辞書 220a から 220d のいずれかを参照して音声認識し、統計的に得られた発話データを基にあらかじめ構成された発話モデルに認識結果をあてはめることにより音声の意味的な言語処理を行う。さらに音声認識言語処理手段 110b は発話モデルと入力された音声との適合度を計算し、適合度に基づいて 1 から n までの重要度を割り付ける。1 フレーズ内で発話モデルとの適合度が最も高い処理単位を重要度 1 とする。適合度の低い部分は冗長部として他の処理単位より大きな値の重要度を割り付け、適合度の高い部分には重要度として小さな値を割り付ける。割り付けた重要度を、入力された音声の処理単位ごとに計時手段 140 および記憶手段 150 へ伝達する。計時手段 140 は音声認識言語処理手段 110b より入力された処理単位ごとの時間長を測定し記憶手段 150 へ伝達する。記憶手段 150 は音声認識言語処理手段 110b より伝達された処理単位ごとに入力音声、重要度および計時手段 140 より伝達された時間長を 1 フレーズ分記憶する。(数 2) の  $i$  に音声認識言語処理手段 110b で割り付けられた重要度をあてはめる。

【0016】 圧縮伸長率決定手段 160b は記憶手段 150 を参照し図 4 の (b) に示すように重要度 1 を割り付けられた処理単位を 1.2 倍に伸長する。図 4 の (c) に示すように、重要度が 4 以上の処理単位を圧縮し重要度が 3 以下の処理単位を伸長する。すなわち、 $\sum T_i$  から  $T_i$  に 1.2 を乗じた値と重要度 4 から n を割り付けられた処理単位について  $T_i$  に 1 未満の  $x_i$  を乗じた値の合計を減じた値と、重要度 2 から 3 を割り付けられた処理単位について  $T_i$  に 1 以上の  $x_i$  を乗じた値の合計が等しくなるよう各  $x_i$  を決定し、処理単位ごとの圧縮あるいは伸長率を制御手段 180 へ伝達する。ただし  $i$  が 1 以外の  $x_i$  は 0.8 以上 1.2 未満であり、 $i$  がより大きい場合により小さい  $x_i$ 、 $i$  がより小さい場合により大きい  $x_i$  を与えるものとする。制御手段 180 は処理単位ごとの圧縮あるいは伸長率を音声速度変換手段 170 に伝達し、入力音声を記憶手段 150 より呼び出して音声速度変換手段 170 に伝達する。音声速度変換手段 170 は制御手段より伝達された圧縮あるいは伸長率に基づいて処理単位ごとに速度変換を行う。

【0017】 (実施例 3) 次に本発明の第 3 の実施例について、図面を参照しながら説明する。

8

【0018】 図 5 は本発明の第 3 の実施例における音声速度変換制御装置の部分を示す構成ブロック図である。図 6 にこの実施例のタイムチャートを示す。第 3 の実施例の構成において第 1 の実施例の構成と同一物または部分については説明を省略し、異なった部分についてのみ説明する。図 1 の音声認識言語処理部 100a が、図 3 において音声認識言語処理部 100c に置き変わり圧縮伸長率決定手段 160a が 160c に変わり、420a ~ n がつけ加わった以外は図 1 と同一な構成である。辞書 420a ~ n はあらかじめ設定された場面に応じた語彙を持つ辞書を格納した記録媒体である。460 は音声認識言語処理手段 110c が参照する辞書 420a ~ n を読み取る辞書読み取り手段である。

【0019】 以上のように構成された第 3 の実施例の音声速度変換制御装置について、その動作を図 5 に従って説明する。まず第 1 の実施例と同様に入力音声は音声認識言語処理部 100c と記憶手段 150 へ伝達される。辞書読み取り手段 460 にあらかじめ場面にに応じてセットされた辞書 420 を辞書読み取り手段 460 を介して参照して、音声認識言語処理手段 110c は音声認識し、統計的に得られた発話データを基にあらかじめ構成された発話モデルに認識結果をあてはめることにより音声の意味的な言語処理を行う。さらに音声認識言語処理手段 110c は発話モデルと入力された音声との適合度を計算し、あらかじめ適合度の低い部分は冗長部として他の処理単位より大きな値の重要度を割り付け、適合度の高い部分には重要度として小さな値を割り付けるよう決められた、各適合度の範囲に対応する 1 から 10 までの重要度を割り付ける。割り付けた重要度を、入力された音声の処理単位ごとに計時手段 140 および記憶手段 150 へ伝達する。計時手段 140 は音声認識言語処理手段 110c より入力された処理単位ごとの時間長を測定し記憶手段 150 へ伝達する。記憶手段 150 は音声認識言語処理手段 110c より伝達された処理単位ごとに入力音声、重要度および計時手段 140 より伝達された時間長を 1 フレーズ分記憶する。(数 3) の  $i$  に音声認識言語処理手段 110c で割り付けられた重要度をあてはめる。圧縮伸長率決定手段 160c は記憶手段 150 を参照し図 6 の (b) に示すように各重要度に対して与えられる圧縮あるいは伸長率  $x_i$  の、重要度ごとの比  $r_i$  をあらかじめ決められた各  $T_i$  に各  $r_i$  を乗じた値をすべての  $i$  について 1 フレーズ分合計した値で  $\sum T_i$  を除した値を  $S$  とする。 $S$  を  $r_i$  に乗じて  $\sum T_i$  に対して  $r_i$  正規化して各  $x_i$  を決定し処理単位ごとの圧縮あるいは伸長率を制御手段 180 へ伝達する。ただし  $i$  がより大きい場合により小さい  $r_i$ 、 $i$  がより小さい場合により大きい  $r_i$  とする。制御手段 180 は処理単位ごとの圧縮あるいは伸長率を音声速度変換手段 170 に伝達し、入力音声を記憶手段 150 より呼び出して音声速度変換手段 170 に伝達する。音声速度変換手段 170 は制御手段より

9

伝達された圧縮あるいは伸長率に基づいて処理単位ごとに速度変換を行う。

【0020】以上のように本発明の3つの実施例によれば、入力音声を音声認識言語処理部100と時間長調整部130へ伝達し、辞書120または場合に応じてスイッチ260により切り替えられる辞書220または辞書読み取り手段460を介して記録媒体に格納された辞書420を参照し、音声認識言語処理手段110において間投詞と主部、述部を構成する品詞を選択する、または統計的な発話データに基づいて構成された発話モデルとの適合度を計算し、意味的な言語処理を行って音声の単位時間ごとの重要度を判断し、で最も重要と判断された処理単位を一定の伸長率で伸長し他の部分に対してはより重要な処理単位には高い伸長率を与え、より重要でない処理単位には低い伸長率あるいは高い圧縮率を与えるあるいは、制御手段180が処理単位ごとに音声速度変換手段170を制御して、重要度に基づき処理単位ごとに音声を圧縮、伸長することにより、重要な語がフレーズの後部に発話された場合も情報が欠落しない圧縮、伸長率を決定する音声速度変換制御装置を構成することができる。

【0021】なお、本発明の実施例1において間投詞に低い重要度を割り付けるとしたが、終助詞、接続助詞、接続詞等の間投詞以外の品詞で情報量の少ない冗長部に低い重要度を割り付けても良い。

【0022】なお、本発明の実施例1において主部、述部を構成する品詞には高い重要度を割り付けるとしたが名詞、動詞、形容詞等の情報量の多い品詞に高い重要度を割り付けても良い。

【0023】なお、本発明の実施例1において間投詞に低い重要度を割り付け主部、述部を構成する品詞には高い重要度を割り付けるとしたが、実施例2のように発話モデルと入力された音声との適合度を計算し、適合度の低い部分は冗長部として低い重要度を割り付け、適合度の高い部分には高い重要度を割り付けてとしても良い。

【0024】なお、本発明の実施例1において音声認識言語処理手段110aは辞書120を参照するとしたが実施例2のように音声認識言語処理手段と複数の辞書との接続をスイッチによって切り替え、接続された辞書を参照するとしても良い。

【0025】なお、本発明の実施例1において音声認識言語処理手段110aは辞書120を参照するとしたが実施例3のように辞書読み取り手段を介して記録媒体に格納された辞書を参照するとしても良い。

【0026】なお、本発明の実施例1において(数1)を用いて圧縮伸長率を決定したが実施例2のように(数2)を用いても良い。

【0027】なお、本発明の実施例1において(数1)を用いて圧縮伸長率を決定したが実施例3のように(数3)を用いても良い。

10

【0028】なお、本発明の実施例2の図2において辞書220はaからdの4種類としたが2種類以上の他の数としても良い。

【0029】なお、本発明の実施例2において、音声認識言語処理手段110bと辞書220a~dの接続の切り換えは切り替え入力によりスイッチ260によって行ったが、これをソフトウェア的に実現しても良い。

【0030】なお、本発明の実施例2において、音声認識言語処理手段110bはあらかじめ場面に応じて切り替え入力により切り替えられたスイッチ260によって接続された辞書220a~dを参照して音声を認識したが、実施例3のように辞書読み取り手段を介して記録媒体に格納された辞書を参照して音声を認識しても良い。

【0031】なお、本発明の実施例2において、音声認識言語処理手段110bはあらかじめ場面に応じて切り替え入力により切り替えられたスイッチ260によって接続された辞書220a~dを参照して音声を認識したが、実施例1のように接続された1つの辞書を参照しても良い。

【0032】なお、本発明の実施例2において、音声認識言語処理手段110bは発話モデルと入力された音声との適合度を計算し、適合度の低い部分は冗長部として低い重要度を割り付け、適合度の高い部分には高い重要度を割り付けるとしたが、実施例1のように間投詞に低い重要度を割り付け主部、述部を構成する品詞には高い重要度を割り付けるとしても良い。

【0033】なお、本発明の実施例2において、(数2)を用いて圧縮伸長率を決定したが実施例1のように(数1)を用いても良い。

【0034】なお、本発明の実施例2において、(数2)を用いて圧縮伸長率を決定したが実施例3のように(数3)を用いても良い。

【0035】なお、本発明の実施例2において、4以上の重要度を割り付けられた処理単位に対して1未満の伸長率すなわち圧縮を与えたが、4以外の2以上20以下の値としても良い。

【0036】なお、本発明の実施例3において、音声認識言語処理手段110cはあらかじめ場面に応じてセットされた辞書420を辞書読み取り手段460を介して参照して音声を認識したが、実施例2のように音声認識言語処理手段と複数の辞書との接続をスイッチによって切り替え、接続された辞書を参照するとしても良い。

【0037】なお、本発明の実施例3において、音声認識言語処理手段110cはあらかじめ場面に応じてセットされた辞書420を辞書読み取り手段460を介して参照して音声を認識したが、実施例1のように接続された1つの辞書を参照しても良い。

【0038】なお、本発明の実施例3において、音声認識言語処理手段110bは発話モデルと入力された音声との適合度を計算し、適合度の低い部分は冗長部として



低い重要度を割付け、適合度の高い部分には高い重要度を割り付けるとしたが、実施例 1 のように間投詞に低い重要度を割付け主部、述部を構成する品詞には高い重要度を割り付けるとしても良い。

【0039】なお、本発明の実施例 3 において、(数 3) を用いて圧縮伸長率を決定したが実施例 1 のように(数 1) を用いても良い。

【0040】なお、本発明の実施例 3 において、(数 3) を用いて圧縮伸長率を決定したが実施例 2 のように(数 2) を用いても良い。

【0041】なお、本発明の実施例 3 において重要度を 1 から 10 までとしたが重要度の最大値は 2 以上の 10 以外の数値としても良い。

【0042】なお、本発明の実施例 1 および 2 において、 $x_1$  は 1.2 としたが 1 より大きく 3 以下の 1.2 以外の値でも良い。

【0043】なお、本発明の実施例 1 および 2 において、 $x_i$  は 0.8 以上としたが 0.3 以上 3 以下の値でも良い。

【0044】

【発明の効果】以上説明したように、本発明によれば、入力された音声認識し、発話中におけるそれぞれの語の重要性を判断するあるいは次に入力される音声の、重要性の高いものは伸長率を大きくし、重要性の低いものは伸長率を小さくするように音声速度変換手段を制御するため、話速の伸長の効果を拡大し、フレーズ後部の情報の欠落なく速度変換することができる音声速度変換制御装置を構成することができ、その実用的効果は大きい。

【図面の簡単な説明】

【図 1】本発明の第 1 の実施例における音声速度変換制

\*御装置の構成ブロック図

【図 2】本発明の第 1 の実施例における音声速度変換制御装置の制御方法を示すタイムチャート

【図 3】本発明の第 2 の実施例における音声速度変換制御装置の構成ブロック図

【図 4】本発明の第 2 の実施例における音声速度変換制御装置の制御方法を示すタイムチャート

【図 5】本発明の第 3 の実施例における音声速度変換制御装置の構成ブロック図

10 【図 6】本発明の第 3 の実施例における音声速度変換制御装置の制御方法を示すタイムチャート

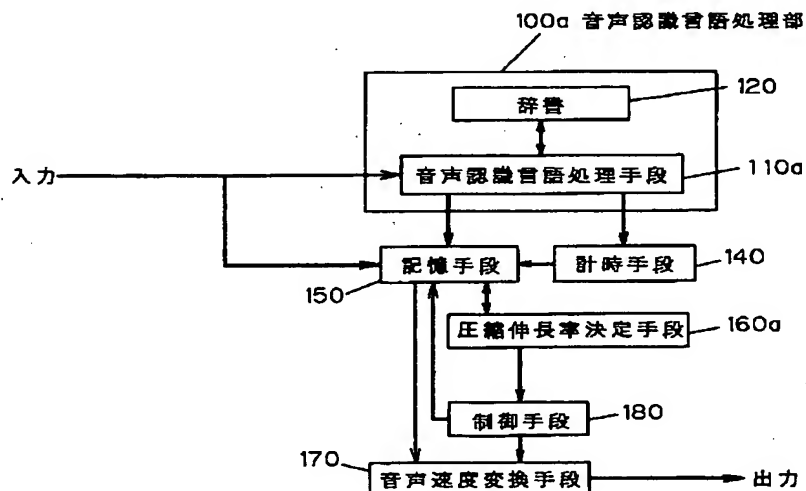
【図 7】従来の音声速度変換制御装置の構成ブロック図

【図 8】従来の音声速度変換制御装置の制御方法を示すタイムチャート

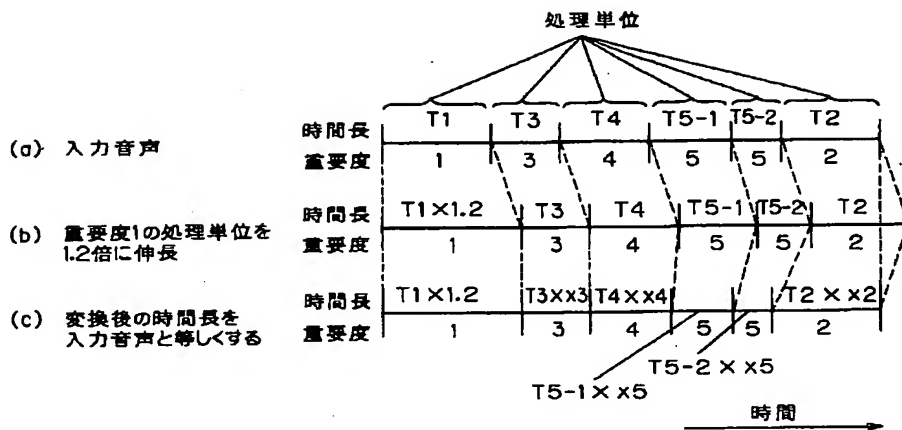
【符号の説明】

- 100 音声認識言語処理部
- 110 音声認識言語処理手段
- 120、220、420 辞書
- 140 計時手段
- 150 記憶手段
- 160 圧縮伸長率決定手段
- 170 音声速度変換手段
- 180 制御手段
- 260 スイッチ
- 460 辞書読み取り手段
- 900 無音区間検出手段
- 910 有声部ピッチ抽出手段
- 920 判定部
- 930 制御手段
- 940 計時手段

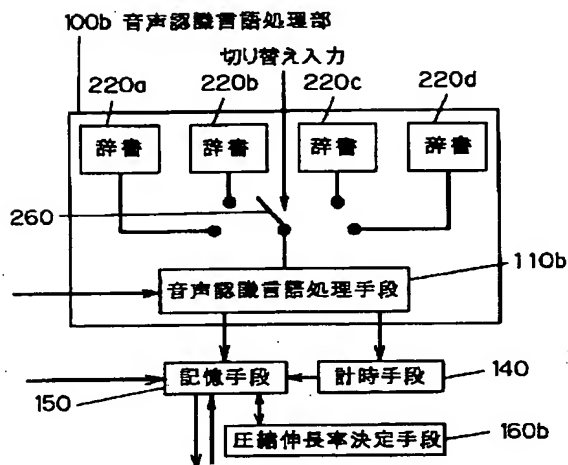
【図 1】



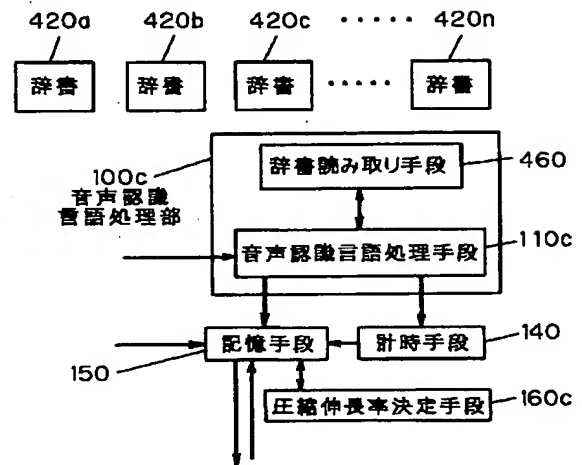
【図 2】



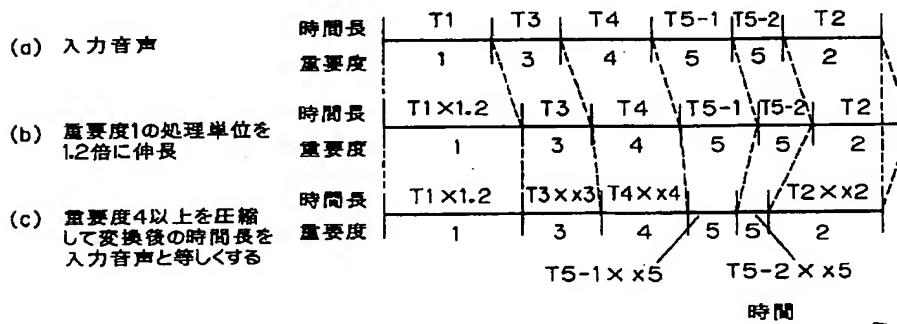
【図 3】



【図 5】

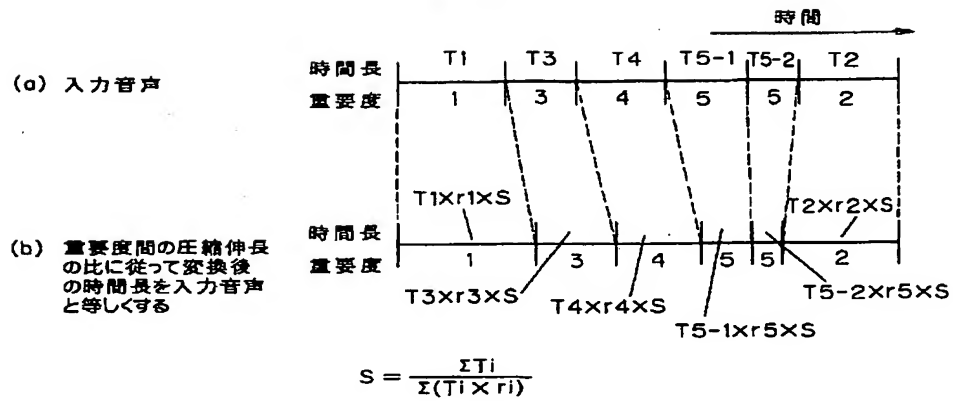


【図 4】



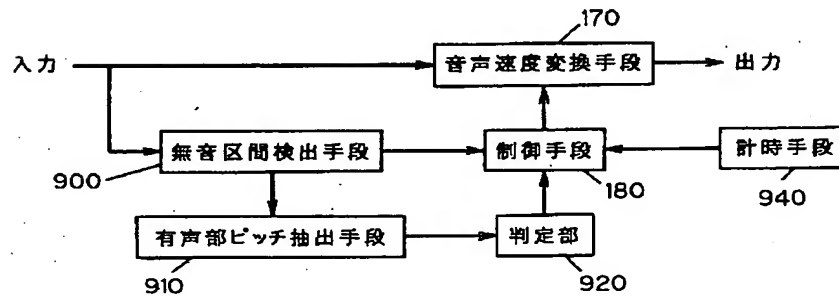


【図 6】



【図 7】

従来例



【図 8】

